

Relativistic Gaussian Functions for Atoms by Fitting Numerical Results with Adaptive Nonlinear Least-Square Algorithm

A. HU,^{1,2} P. OTTO,¹ J. LADIK¹

¹Chair for Theoretical Chemistry and Laboratory of the National Foundation for Cancer Research, Friedrich-Alexander University Erlangen–Nürnberg, D-91058 Erlangen, Egerlandstrasse 3, Germany

²Institute of Solid State Physics, Siping Normal College, Siping, People's Republic of China

Received 11 March 1998; accepted 5 November 1998

ABSTRACT: The relativistic Gaussian basis set of Hg has been obtained by fitting the numerical relativistic atomic radial wave functions with the adaptive nonlinear least-square algorithm combined with the subset selection method. This fitting procedure and fitted results are presented. From this basis, in a further several-step procedure, we generated a new basis that gave reasonably good results for Hg, HgO, and a Se atom chain, respectively. The original fitted basis did not work, because the resulting overlap matrix was not positive definite. © 1999 John Wiley & Sons, Inc. *J Comput Chem* 20: 655–664, 1999

Keywords: relativistic Gaussian basis set; adaptive non-linear fitting; relativistic Hg atom calculation

Introduction

The development of relativistic computational methods that are potentially useful for chemical and biological (iron in hemin, heavy trace elements etc.) applications follows traditional approaches used in quantum chemistry.^{1–6} It is a common practice to use the method of expansion in a finite set of basis functions, as first proposed

Correspondence to: J. Ladik; e-mail: ladik@pctc.chemie.uni-erlangen.de

Contract/grant sponsor: Deutsche Forschungsgemeinschaft; contract/grant number La 371/27-2

for nonrelativistic quantum mechanics by Hall⁷ and Roothaan.⁸ In this way, the problem of solving differential or integrodifferential equations is converted into the much simpler algebraic problem of determination of the eigenvalues of hermitian matrices. Relativistic molecular programs applying this LCAO approach with Cartesian Gaussian basis functions have been used, because, in this way, all types of integrals can be calculated analytically. Therefore, the question arises as to how to construct basis sets for relativistic calculations.

In general, there are several ways to construct relativistic Gaussian basis sets for four-component spinors on the basis of relativistic total energy.

One can reoptimize nonrelativistic basis sets for functions of large components and thus small components are generated by kinetic and atomic balance conditions.⁹

Further applying the integral discretization technique to the generator coordinate version of the Hartree–Fock method^{10,11} a universal (for each valid atom) Gaussian basis set was developed.^{12,13} Clementi and coworkers, starting from the numerical finite differences (GRASP2) method,¹⁴ developed the so-called geometrical basis set.¹⁵ Both methods were applied for different atoms, using a finite nucleus with a Gaussian distribution of positive charge for the universal basis set and with a uniform charge distribution (spherical ball)^{16,17} for the geometrical basis set.

There are three recent studies that also should be mentioned. In one of these, DaSilva and coworkers¹⁸ improved the universal basis set by optimizing the orbital exponents for each atom separately, using the Dirac–Fock (DF) program (adapted Gaussian basis). In this way, again taking a finite nucleus, they obtained somewhat better results than any other method has achieved until now. In a calculation of the excited states of the Hg atom after a modified Douglas–Kroll transformation, Gleichmann and Hess¹⁹ obtained good agreement with experiment²⁰ with their two-component wave functions after introducing a multireference singles and doubles CI to take correlation into account. Unfortunately, they did not give any value for the DF ground state energy of the Hg atom. Finally, Dyall and Faegri,²¹ in a very interesting study, generally investigated the problem of four component spinors constructed from Gaussians employing analytic gradients and a scaled Newton–Raphson procedure.²¹ The method has been implemented for closed-shell systems as a subprogram of the first version of the GRASP package.²² They concluded that Gaussians having the same l values, but different j ($= l - 1/2, l + 1/2$) values, should have different orbital exponents. In other words, the exponents should be the same for a given n and j , like $s_{1/2}$, $p_{1/2}$ or $p_{3/2}$, $d_{3/2}$. They argued that it would be still better to optimize a set of exponents for each $\kappa = (j + 1/2, -(j + 1/2))$. Also, using a finite nucleus they performed a series of calculations on the Rn atom. Their results verified theoretical conclusions.

One can also generate basis sets by fitting the numerical output of a numerical MCDP program, like Desclaux's program,²³ with a linear combination of Gaussians.²⁴ In this article, we present a fitting procedure using the Hg atom as an exam-

ple. For this, the adaptive nonlinear least-square algorithm of Dennis²⁵ and the statistical subset selection²⁶ have been used for the computation of the set of Gaussian radial basis functions. In the next section, we present the theory of the adaptive nonlinear least-square algorithm by comparing the Gauss–Newton method with the usual Newton method. The backward elimination in the statistics subset selection is introduced into our procedure to obtain more accurate fittings. The resulting basis set for the Hg atom is given in the third section, together with some calculated results for the Hg atom by using this basis set.

Theory

The relativistic atomic radial wave functions that can be calculated by the MCDP-94 code of Desclaux²³ are more complicated than the nonrelativistic ones. For example, the relativistic hydrogen-like functions consist essentially of a product of a confluent hypergeometric function with a nonintegral power of r . For the fitting of these functions with the Gaussian basis functions, some sophisticated methods are needed. Here, we setup a fitting procedure by using the adaptive nonlinear least-square algorithm, combined with the subset selection method.

The radial parts of the Gaussian basis functions, which were used for the fitting have the form:

$$y = y(r, c_\alpha, \alpha) = c_\alpha r^n e^{-\alpha r^2} \quad (1)$$

Here, α and c_α are the exponent and the LCAO coefficient, respectively, and n defines the possible l -values in a given AO. The nonlinear fitting procedure has been carried out on a set of M unknown parameters, a_j , $j = 1, \dots, M$, which consist of the exponents α and the LCAO coefficients c_α . We first define a χ^2 merit function and then determine the best fitting parameters by its minimization. With nonlinear dependence, however, the minimization must proceed iteratively. Given trial values for the parameters, we can develop a procedure that improves the trial solutions on the basis of Dennis's adaptive nonlinear least-square algorithm.²⁵ The χ^2 merit function is defined as:

$$\chi^2(\{a_j\}_{j=1}^M) = \frac{1}{2} \sum_{i=1}^m \left[y_i - y(x_i; \{a_j\}_{j=1}^M) \right]^2 \quad (2)$$

where the residuals, $r_i(\{a_j\}_{j=1}^M) = y_i - y(x_i; \{a_j\}_{j=1}^M)$, are functions of the parameters, $\{a_j\}_{j=1}^M$, as is the

merit function, χ^2 . Finally, m is the number of grid points used in the fitting procedure. The sum of components of its gradient with respect to parameters $\{a_j\}_{j=1}^M$ is:

$$\sum_{t=1}^M \nabla_{a_t} \chi^2(\{a_j\}_{j=1}^M) = \sum_t \sum_j r_i(\{a_j\}_{j=1}^M) \frac{\partial r_i(\{a_j\}_{j=1}^M)}{\partial a_t} \quad (3)$$

and its Hessian matrix is:

$$\frac{\partial^2 \chi^2(\{a_j\}_{j=1}^M)}{\partial a_s \partial a_t} = 2 \sum_{i=1}^M \left[\frac{\partial r_i(\{a_j\}_{j=1}^M)}{\partial a_s} \frac{\partial r_i(\{a_j\}_{j=1}^M)}{\partial a_t} + r_i(\{a_j\}_{j=1}^M) \frac{\partial^2 r_i(\{a_j\}_{j=1}^M)}{\partial a_s \partial a_t} \right] \quad (4)$$

In fact, it is often possible to ignore the second term of the Hessian matrix:

$$\sum_i r_i(\{a_j\}_{j=1}^M) \frac{\partial^2 r_i(\{a_j\}_{j=1}^M)}{\partial a_s \partial a_t} \quad (5)$$

on the basis that if the non-zero residuals are such that they do not reinforce their nonlinearity, the term

$$\sum_i \frac{\partial r_i(\{a_j\}_{j=1}^M)}{\partial a_s} \frac{\partial r_i(\{a_j\}_{j=1}^M)}{\partial a_t} \quad (6)$$

is a sufficiently good approximation. In the resulting Gauss-Newton method, the "Newton k th step" from the a_j^k to $a_j^k + \delta a_j^k$ with the change, δa_j^k , is defined by the linear system of equations:

$$\sum_j A_{jq} \delta a_q = B_j \quad (q = 1, \dots, M) \quad (7)$$

where A_{jq} and B_j are:

$$A_{jq} = \sum_l \frac{\partial r_l(\{a_j^k\}_{j=1}^M)}{\partial a_j} \frac{\partial r_l(\{a_j^k\}_{j=1}^M)}{\partial a_q} \quad (8)$$

$$B_j = - \sum_l r_l(\{a_j^k\}_{j=1}^M) \frac{\partial r_l(\{a_j^k\}_{j=1}^M)}{\partial a_j} \quad (9)$$

One can compare this quadratic Taylor expansion of the least-square merit function, $\chi^2(\{a_j\}_{j=1}^M)$,

around the parameter values, $\{a_j^k\}_{j=1}^M$:

$$\begin{aligned} \chi^2\{a_j\}_{j=1}^M &= \chi^2(\{a_j^k\}_{j=1}^M) \\ &+ 2 \sum_t \sum_i r_i(\{a_j^k\}_{j=1}^M) \frac{\partial r_i(\{a_j^k\}_{j=1}^M)}{\partial a_t} (a_t - a_t^k) \\ &+ \sum_{t,s} \sum_i \left[\frac{\partial r_i(\{a_j^k\}_{j=1}^M)}{\partial a_s} \frac{\partial r_i(\{a_j^k\}_{j=1}^M)}{\partial a_t} \right. \\ &\left. + r_i(\{a_j^k\}_{j=1}^M) \frac{\partial^2 r_i(\{a_j^k\}_{j=1}^M)}{\partial a_s \partial a_t} \right] (a_s - a_s^k)(a_t - a_t^k) \end{aligned} \quad (10)$$

with the Gauss-Newton model [eq. (4)].

In the nonlinear least-square algorithm of Dennis, the method will be used in an adaptive manner through the comparison of the results of the Gauss-Newton and the usual Newton methods. A trust region for minimization has been implemented in this procedure. This means that around the parameter values $\{a_j^k\}_{j=1}^M$ of the k th step, there is a local quadratic expression [eq. (10)] of the merit function, $\chi^2(\{a_j\}_{j=1}^M)$, and a region is estimated in which the local quadratic model is trusted to represent the merit function. In the next step, the parameters $\{a_j^{k+1}\}_{j=1}^M$ are chosen to approximately minimize this quadratic expression in this region or to minimize as an approximation of the region itself. In either case, the information gained about the merit function $\chi^2(\{a_j\}_{j=1}^M)$ via the parameters $\{a_j^{k+1}\}_{j=1}^M$ is then used to update the model and also to update the size or shape of the trust region.

In our fitting procedure, we used the just described algorithm. Also, for subset selection in statistics,²⁷ we define our radial Gaussian function as:

$$y(x; a_1, a_2, a_3) = a_1 r^n e^{-a_2(r-a_3)^2} \quad (11)$$

with center a_3 ; a_1 is again an LCAO coefficient, and a_2 an exponent. In the fitting procedure, we treated the origin of the AOs as a variational parameter to have a more flexible function and to avoid obtaining local minima. Thereafter, we set the a_3 parameters equal with the position of the nucleus²⁸ (see later). For subset selection, there are two methods: forward selection and backward elimination. The latter starts with full sets and

removes one parameter at a time—the one that least increases the merit function—until the chosen criterion is met. In our procedure, we used backward elimination to remove the center of the Gaussian functions, also in an adaptive manner (i.e., to make equal to the position of the nucleus). Then, from the results obtained for the merit function, $\chi^2(\{a_j\}_{j=1}^M)$, it is determined which a_3 (center) parameter should be removed. Doing so, finally, we removed all a_3 parameters (by setting them equal with the position of the nucleus). It is of interest to compare this method with the standard procedure for optimizing a neural network. The latter involves optimization, by gradient descent, of a nonlinear, sum-of-squares error surface in a high-dimensional space defined by network parameters.²⁸ Tests have shown that, with the help of backward elimination, the merit function can be decreased by one order of magnitude.

Fitting Relativistic Atomic Orbitals with Gaussians and Generation of Basis Set for Hg Atom

Using the adaptive nonlinear least-square algorithm and back elimination method, we have fitted the relativistic numerical radial wave functions with Gaussians for state $6s^2$ and state $6s^16p^1$ of the Hg atom. For the $6s^16p^1$ state, the $6p_{\frac{1}{2}}$ and $6p_{\frac{3}{2}}$ functions can be obtained to fit the excited state wave functions. Also, the wave functions of the small components have been fitted with this procedure so that one can verify the working of this algorithm although the basis functions of the small components can be generated by kinetic and atomic balance conditions.⁹ Generally speaking, fitting a radial wave function belonging to higher principal quantum numbers is more difficult than with lower n values, because the former has more nodes than the latter ones and therefore needs more Gaussians for the fit. Fitting small components is also more difficult than fitting large components, because small components are *much* smaller. In Figure 1, we show only the numerically obtained results for the large component of $6p_{\frac{3}{2}}$. One can see that, in this case also, the fit is quite good, even when close to the nucleus (see inset).

The fitted exponents and coefficients are shown in Tables I to III for states $1s_{\frac{1}{2}}$, $6s_{\frac{1}{2}}$, and $6p_{\frac{3}{2}}$, respectively. As the exponents are obtained with the help of a fitting procedure, it is very difficult, probably impossible, to interpret the occurrence of

the closely spaced values of the Gaussian exponents (see Tables II and III).

The corresponding number of gridpoints and values of the merit functions are also given in the tables. Despite an intensive literature search it was not possible to find a reference on how the gridpoints are chosen in the Desclaux program. Therefore, we studied the program code with regard to this question. A maximum number of gridpoints was chosen (800 for our calculation) for the radial wave functions. The actual number of grid points for the output of the radial wave functions was determined according to the range between two minimum values of the radial wave functions. For the minimum near the origin, and for long distance (approaching infinity), values $1.0\text{E-}05$ and $1.0\text{E-}15$ were taken. This means that the range for the output radial functions was determined by the convergence of the radial function when the radial variable went to the origin and to infinity, respectively. The gridpoints were then chosen within the space between the two minima (which change in the course of the self-consistent-field procedure). However, the detailed algorithms and how this step was done in the program, as well as determination of the respective weighting factors, could not be determined by studying the source code of the Desclaux program.²³ In general, the resulting number of grid points of the output radial function must be less than that of the given maximum. On the basis of this concept it is also understandable that the number of gridpoints was different for different radial wave functions (see Tables II and III). It can be seen that the large and corresponding small components of the $1s_{\frac{1}{2}}$, $6s_{\frac{1}{2}}$, $6p_{\frac{3}{2}}$ radial wave functions of Hg were very well fitted by 8 and 12 Gaussian functions, respectively. In this procedure, the initial guess is very important because a good guess can make convergence much faster. In our experience, it is better to take the fitted exponents and coefficients of an orbital with a higher symmetry as a guess for the orbital with a lower symmetry. For example, we took the fitted exponents and coefficients of the $6s_{\frac{1}{2}}$ wave function as a guess for the $6p_{\frac{3}{2}}$ wave function. Doing so, allows for convergence with a desired accuracy.

As starting point for developing our crystal orbital program²⁴ we used the MOLFDIR program of Nieuwpoort et al.⁴ This program, which accepts as input the universal, the geometrical, and other basis sets, did not accept our fitted basis functions numerical results. The reason for this was that the overlap matrix of our basis functions was not

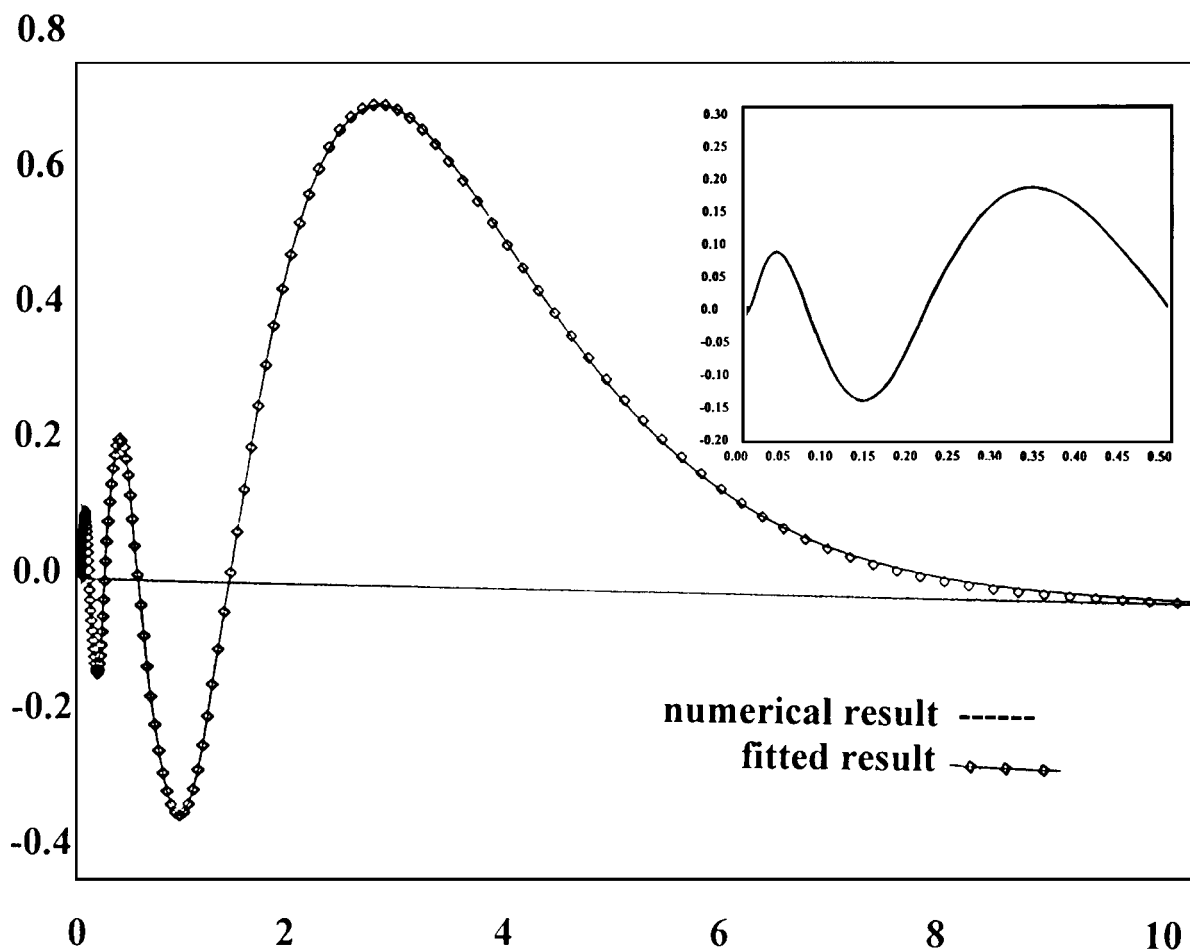


FIGURE 1. The large component of the $6p_{3/2}$ orbital of Hg calculated with the Desclaux program and fitted by 12 Gaussians.

TABLE I.
Desclaux Numerical $1s_{\frac{1}{2}}$ Wave Function of Hg Fitted by Gaussians at 327 Gridpoints.

1s (large component)			1s (small component)		
Gaussian	Coefficient	Orbital Exponent	Gaussian	Coefficient	Orbital Exponent
1	-0.306169	1.109248E + 8	1	0.140901	4.298753E + 7
2	-0.694475	6.712076E + 6	2	0.251142	4.032080E + 6
3	-1.378095	7.079587E + 5	3	0.460584	5.268762E + 5
4	-2.390282	1.035129E + 5	4	0.756725	8.849955E + 4
5	-2.893565	1.909712E + 4	5	0.873576	1.804017E + 4
6	4.764159	1.369554E + 3	6	-1.443900	1.440318E + 3
7	2.862308	5.061444E + 2	7	-0.980069	5.478350E + 2
8	0.095866	1.723128E + 2	8	-0.059915	2.130288E + 1
Merit function $\chi^2 = 7.7790\text{E} - 3$			Merit function $\chi^2 = 7.1894\text{E} - 4$		

TABLE II. Desclaux Numerical $6s_{\frac{1}{2}}$ Function of Hg Fitted by Gaussians at 379 Gridpoints.

6s (large component)			6s (small component)		
Gaussian	Coefficient	Orbital Exponent	Gaussian	Coefficient	Orbital Exponent
1	30.246894	8.931923E + 5	1	30.259452	1.242197E + 6
2	-21.538725	8.946396E + 5	2	-21.504495	1.261271E + 6
3	-8.662568	8.916101E + 5	3	-8.767912	1.196133E + 6
4	8.047063	3.753259E + 3	4	-7.992034	2.541624E + 4
5	7.916381	3.839689E + 3	5	7.971148	2.541624E + 4
6	4.990569	1.750888E + 2	6	4.615844	9.861298E + 1
7	-4.342002	2.019401E + 2	7	-4.693215	9.781138E + 1
8	-1.982832	1.658613E + 1	8	-2.180269	1.111150E + 1
9	-5.420945	6.790293E - 1	9	-4.387794	1.786975E + 0
10	2.522003	4.888551E + 0	10	2.244866	1.124658E + 1
11	3.333380	6.790260E - 1	11	4.366366	1.775717E + 0
12	0.980695	5.766937E - 2	12	0.068524	5.848402E + 2
Merit function $\chi^2 = 2.2716E - 4$			Merit function is $\chi^2 = 6.4009E - 5$		

positive definite, which turned up when we tried to solve the generalized eigenvalue problem. Therefore, we had to generate a new several-step procedure to obtain a new basis that is acceptable for the MOLFDIR program.

We emphasize that if one desires to go to molecules and particularly infinite chain calculations, one cannot use such large basis sets as for atoms. We feel that our procedure provides an acceptable compromise (see below Table VI) be-

tween accuracy and applicability to larger systems. The reason for this is that, by starting from basis functions well fitted to numerical results, the sensitive core region is better described than with other basis sets of the same size generated in different ways.

Finally, we point out that the MOLFDIR program uses a double group for a unitary transformation, which gives j -dependence of the basis functions belonging to the same n value. In this

TABLE III. Desclaux Numerical $6s_{\frac{3}{2}}$ Function of Hg Fitted by Gaussians at 357 Gridpoints.

$6p_{\frac{3}{2}}$ (large component)			$6p_{\frac{3}{2}}$ (small component)		
Gaussian	Coefficient	Orbital Exponent	Gaussian	Coefficient	Orbital Exponent
1	31.789144	3.796924E + 5	1	30.114778	5.934452E + 3
2	-22.307719	3.824079E + 5	2	-21.837283	5.934455E + 3
3	-9.483564	3.732756E + 5	3	-8.272108	5.934385E + 3
4	-8.070823	4.708682E + 3	4	-7.968713	8.674060E + 2
5	8.027984	4.708682E + 3	5	7.994999	8.674064E + 2
6	6.255017	8.773998E + 2	6	4.617927	1.074994E + 2
7	-6.426647	8.773998E + 2	7	-4.683697	1.069671E + 2
8	-1.597680	1.689144E + 1	8	-2.200682	8.485205E + 0
9	-2.319558	4.190053E - 1	9	-4.809537	2.022224E + 0
10	2.040096	3.374167E + 0	10	2.219187	8.690491E + 0
11	1.264094	6.607248E - 2	11	3.960101	1.949589E + 0
12	0.829637	9.054653E + 1	12	0.865000	2.414373E + 0
Merit function $\chi^2 = 1.2817E - 4$			Merit function $\chi^2 = 6.9109E - 5$		

way, the linear dependencies caused by the same orbital exponents for a given n value do not occur.

After ordering the fitted exponents of the s -type Gaussian functions according to increasing values we determined the lowest and highest exponents of the p , d , and f basis functions from the universal and geometrical basis sets. For this purpose we chose those exponents of the s functions for the p , d , and f orbitals that were nearest to the predicted limiting values given by the universal basis set of Malli and DaSilva^{12,13} and by the geometrical basis set of Clementi et al.,¹⁵ respectively. In this way, we obtain p , d , and f basis functions, which are subsets of s orbitals.

Having established the preliminary total basis set for the large component we reoptimized the exponents of the s orbitals with respect to the atomic energy following the SCF procedure of Clementi et al.¹⁴ The exponents of the 22s final basis functions for the large component of the Hg

atom are listed in Table IV together with those 22 exponents of the universal basis set closest to them in the case of the Hg atom.

The basis set developed in the previous procedure is uncontracted. To obtain the contraction scheme we used the LCAO coefficients from the previous step.

The basis set for the small components could have been generated in the same way as just described, but this would have resulted in extremely time-consuming calculations. Therefore, we compromised and used the atomic and kinetic balance relations to obtain them. The whole contraction scheme is given in Table IV.

Results and Discussion

The one-electron energies from our basis, generated as described earlier, and the results of numer-

TABLE IV.
Basis Sets for Hg Atom and Contractions.

Exponents	s	p	d	f
39147653.317000 ^a	27508345.793637 ^b	×		
6048148.218000	7802853.046713	×		
1441981.718000	1178791.123851	×		
374566.768000	334368.848683	×		
133241.967000	178082.107320	×	×	
44221.682900	50513.706789	×	×	
18301.990800	14328.416324	×	×	
6988.521800	7631.197056	×	×	×
3030.079100	2164.619772	×	×	×
1311.572200	1152.858743	×	×	×
564.047810	614.003114	×	×	×
250.879870	174.164456	×	×	×
118.636540	92.758561	×	×	×
57.686319	49.402449	×	×	×
28.623031	26.311339	×	×	×
13.993355	14.013204	×	×	×
6.715883	7.463317	×	×	×
3.160740	3.974902	×	×	×
1.324000	1.127497	×	×	×
0.517815	0.600496	×	×	×
0.176426	0.170333	×	×	×
0.046904	0.048316	×	×	
Large component		Small component		
Primitive	Contracted	Primitive	Contracted	
22s18p14d8f	8s10p8d3f	18s22p18d14f8g	12s18p14d8f	

^aObtained by the described procedure.

^bSelected from the universal basis set of 40s functions¹² as nearest values to the 22 values originated by our procedure.

TABLE V.
Relativistic Orbital Energies (a.u.) of Hg Atom.

Orbital	Symmetry Rep.	Present Basis	Numerical ²³
1s _{1/2}	e _{2g} ¹	−3074.243493	−3074.233600
2s _{1/2}	e _{2g} ¹	−550.259197	−550.251980
2p _{1/2}	e _{2u} ¹	−526.856856	−526.854520
2p _{3/2}	e _{2u} ¹	−455.164535	−455.156470
3s _{1/2}	e _{2g} ¹	−133.117836	−133.113330
3p _{1/2}	e _{2u} ¹	−122.629049	−122.638860
3p _{3/2}	e _{2u} ¹	−106.549618	−106.545040
3d _{3/2}	e _{2g} ¹	−89.437071	−89.436684
3d _{5/2}	e _{2g} ¹	−86.024389	−86.020000
4s _{1/2}	e _{2g} ¹	−30.651514	−30.648355
4p _{1/2}	e _{2u} ¹	−26.123559	−26.124082
4p _{3/2}	e _{2u} ¹	−22.191764	−22.188524
4d _{3/2}	e _{2g} ¹	−14.798791	−14.796681
4d _{5/2}	e _{2g} ¹	−14.055761	−14.052491
4f _{5/2}	e _{2u} ¹	−4.475476	−4.472838
4f _{7/2}	e _{2u} ¹	−4.314561	−4.311635
5s _{1/2}	e _{2g} ¹	−5.105832	−5.103105
5p _{1/2}	e _{2u} ¹	−3.539723	−3.537924
5p _{3/2}	e _{2u} ¹	−2.844605	−2.841956
5d _{3/2}	e _{2g} ¹	−0.652533	−0.650063
5d _{5/2}	e _{2g} ¹	−0.577178	−0.574638
6s _{1/2}	e _{2g} ¹	−0.328782	−0.328042
Total energy		−19648.835901	−19648.877845
Gaunt interaction		25.351139	25.353550

ical calculations of the ¹S ground state of Hg²³ are shown in Table V. There is better agreement in the valence region than in the core region.

It should be pointed out that the core levels, especially the innermost spinors, are influenced by the nuclear model used. In the numerical calculations, a uniform charge sphere was used and, in the finite basis set calculations, a Gaussian distribution was employed with the exponent 0.1401178891 × 10⁹ for Hg.

The relativistic Dirac–Coulomb total energy of the ¹S ground state of the Hg atom is −19648.835901 a.u. using the basis set generated by the procedure just described (see Table VI) compared with the numerical energy of −19648.877854 a.u., which was calculated using the MCDP-94 code.

Table VI gives the Dirac–Hartree–Fock total energies of the ground state of the Hg atom using the universal basis set (again, with finite nucleus

having a Gaussian charge distribution with an (38s38p25d19f) basis.^{12,13} We also included in Table VI the results of applying the geometrical basis¹⁴ and those of numerical calculation with the GRASP2 package. Finally, we attempted also to use the universal basis with the same number of functions (22s18p14d8f) as was done in our procedure (point 4, in parentheses in Table VI). Here, we were faced with the problem (since there is no specific way) of how to choose 22 Gaussian s functions out of 40. We took those 22 functions that had exponents closest to one of our 22s functions. As one can see from the results the resulting basis is unbalanced providing an overly high (too small negative) total energy.

One can further see from Table VI that the very sophisticated GRASP2 numerical procedure has provided the best result, followed by the very large universal basis calculation.^{12,13} The next best calculation with nearly the same number is provided by Desclaux’s numerical procedure.²³ The geometrical basis gives a somewhat less satisfactory result¹⁴ and our basis which was constructed to perform also calculations on molecules³⁰ and polymers³¹ gives a nonnegligibly worse result. On the other hand our basis is only moderately large, and the difference between the best Hg *E*_{tot}^{rel} calculation of the GRASP2 package and ours is only 0.044 a.u. (= 1.04 eV). The attempt to use the universal basis set with the same number of basis functions as we did had failed (*E*_{tot}^{rel} is, by 11.31 a.u. [= 307.6 eV], more positive than our result).

Conclusion

Determination of basis functions in the just described way inevitably represents a compromise between the accuracy and economy of any calculation, which is needed if one intends also to use the basis for molecules and polymers. The question how many Gaussian basis functions are needed for a fitting with a desired accuracy have not so far been dealt with in this kind of fitting procedure. However, the forward selection method can be used to do this because the method starts with a few Gaussians as a subset to which one Gaussian at a time is added—the one that reduces the merit function χ^2 —until some chosen criterion is met. In this way, one can determine the minimum number of Gaussian functions for fitting with a prescribed accuracy. We plan to investigate this fitting procedure for heavy atoms.

TABLE VI.
Relativistic Total Energies of Hg Atom Using Different Basis Sets.

		$E_{\text{total}}^{\text{DHF}}$ (a.u.)
1.	Present calculation ([22s18p14d8f] basis; finite nucleus with a Gaussian charge distribution)	− 19648.8359 (25.3511) ^a
2.	Numerical calculation (Desclaux ²³ finite nucleus with uniformly distributed charge)	− 19648.8778 (25.3536) ^a
3.	Universal basis set ([38s38p25d19f] ^{12, 13} ; finite nucleus with a Gaussian charge distribution ^{16, 17})	− 19648.8789
4.	Universal basis set ([22s18p14d8f]; finite nucleus with a Gaussian charge distribution) ^b	− 19637.5219 (25.2789) ^a
5.	Geometrical basis set ([33s,28p] ¹⁵ ; finite nucleus, uniform charge distribution) ^c	− 19648.8720
6.	Numerical calculation with the GRASP2 package (finite nucleus ¹⁶)	− 19648.8800

^aIn parentheses are the energies of the Gaunt term.^bThe same number of basis functions as in our calculation was selected from the universal basis set see eq. (3), taking, as exponents, those values of the universal basis that lie nearest to the ones of our basis.^cThe number of *d* and *f* functions is not given in ref. 15.

Recently, we have used our basis set generated by the procedure described to perform a calculation on the HgO molecule.³⁰ The total energy belonging to the Dirac–Coulomb Hamiltonian, the Gaunt correction, as well as the binding energy, have been obtained. The obtained binding energy of 39.7 kcal/mol is more than two thirds of the experimental value of 53.8 kcal/mol,³² which is unusually good for Dirac–Hartree–Fock calculations without correlation. The optimized bond length was 2.00, which is in good agreement with the experimental value of 2.03 obtained from X-ray diffraction.³³

The relatively good results obtained with our comparatively small basis are most probably due to the fact that we started its generation by fitting (by Gaussians) the numerical solutions of the DHF equations for the Hg atom and therefore described the core region better than other small basis sets constructed in a different way. We could probably take still more advantage of the fitting procedure if we: (1) used a somewhat larger number of Gaussians; and (2) introduced contractions of the same type of functions (possibly according to *j* or *κ*) already in the fitting procedure (hopefully

avoiding, in this way, the difficulty of negative eigenvalues of the overlap matrix).

Acknowledgments

The authors express their gratitude to Dr. L. Vischer for his very useful advice.

References

- Kim, Y. K. *Phys Rev* 1967, 154, 17; 1967, 159, 190.
- Malli, G. L. ed. *Relativistic Effects in Atoms, Molecules and Solids*, NATO ASI Series; Plenum Press: London, 1981.
- Dyall, K. G.; Taylor, P. R.; Faegri, K., Jr., Partridge, H. *J Chem Phys* 1991, 95, 2583.
- Visser, O.; Visscher, L.; Aerts, P. J. C.; Nieuwpoort, W. C. *J Chem Phys* 1992, 96, 2910.
- Kagawa, T. *Phys Rev* 1975, A12, 2245.
- Pisani, L.; Clementi, E. *J Chem Phys* 1988, 89, 385.
- Hall, G. G. *Proc R Soc* 1951, A205, 541.
- Roothaan, C. C. J. *Rev Mod Phys* 1951, 23, 69.
- Visscher, L.; Aerts, P. J. C.; Visser, O.; Nieuwpoort, W. C. *Int J Quant Chem Quant Chem Symp* 1991, 25, 131.
- Mohallem, J. R.; Dreizler, R. M.; Trsic, M. *Int J Quant Chem* 1986, S29, 45.

11. DaSilva, A. B. F.; DaCosta, H. E. M.; Trsic, M. *Mol Phys* 1989, 68, 433.
12. Malli, J. G. L.; DaSilva, A. B. F. *Phys Rev* 1993, A47, 143.
13. Malli, J. G. L.; DaSilva, A. B. F.; Ishikawa, Y. *Chem Phys* 1994, 101, 6829.
14. Gianolio, L.; Pavoni, R.; Clementi, E. *Gazz Chim Ital* 1978, 180, 181.
15. Mohanty, A. K.; Parpia, F. A.; Clementi, E. In: Clementi, E. ed. *MOTECC-91*; Escom: Leiden, 1991; p 167.
16. Ishikawa, Y.; Quinley, H. M. *Int J Quant Chem* 1987, S21, 523.
17. Werner, G.; Wheeler, J. A. *Phys Rev* 1958, 109, 126.
18. Jorge, F. E.; Bobbie, T. B.; DaSilva, A. B. F. *Chem Phys Lett* 1996, 263, 775.
19. Gleichmann, M.; Hess, B. A. *Chem Phys Lett* 1994, 227, 229.
20. Moore, C. E. *Atomic Energy Levels*. NSRDS-MBS Circular 467, Vol. 3; NSRDS-MBS: Washington, DC, 1971.
21. Dyall, K. G.; Faegri, K., Jr. *Theor Chim Acta* 1996, 94, 39.
22. Dyall, K. G.; Grant, I. P.; Johnson, C. T.; Perpia, F. A.; Plummer, E. P. *Comput Phys Commun* 1989, 55, 425.
23. Desclaux, J. P. *Comput Phys Commun* 1973, 9, 87; Desclaux, J. P. *Chem Phys* 1978, 34, 261, and references therein.
24. Ladik, J. *Theochem* 1997, 391, 1; Ladik, J.; Čížek, J.; Mukherjee, K. In: Malli, G. L., ed. *Relativistic Effects in Atoms, Molecules and Solids*, Plenum Press: London, 1983; p. 305.
25. Dennis, J. E. *ACM Trans Math* 1981, 18, 349.
26. Orr, M. J. L. *Neural Comput* 1995, 7, 606.
27. Rawlings, J. O. *Applied Regression Analysis*; Wadsworth & Brooks/Cole: Pacific Grove, CA, 1988.
28. Orr, M. J. L. <http://www.cns.ed.ac.uk/people/mark.html>.
29. Herts, J.; Knough, A.; Palmer, R. G. *Introduction to the Theory of Neural Computation*; Addison Wesley: Redwood City, CA, 1991.
30. Hu, A.; Otto, P.; Ladik, J. *J Chem Phys* In press.
31. Hu, A.; Otto, P.; Ladik, J. *Chem Phys Lett* In press.
32. Grad, M.; Hirswald, W. *Buns Phys Chem* 1982, 86, 899.
33. Scavnicar, S.; Grdenic, D. *Acta Crystallogr* 1955, 8, 275.